DOCUMENT RESUME

ED 170 018 JC 790 327

AUTHOR Hartnett, Carolyn G.
TITLE Measuring Writing Skills.
SPONS AGENCY Advanced Institutional Development Program Two-Year College Consortium.
PUB DATE 78
NOTE 67p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Cognitive Processes; Comparative Analysis; Composition Skills (Literary); Computer Oriented Programs; Computer Programs; Diagnostic Tests; Measurement Instruments; *Measurement Techniques; *Norm Referenced Tests; *Postsecondary Education; Sentence Combining; Standardized Tests; *Student Testing; *Writing; *Writing Skills

ABSTRACT

Measuring student writing skills can be done holistically, by ranking compositions without enumerating their linguistic, rhetorical, or information features, or atomistically, by viewing compositions as collections of these features. A variety of holistic approaches, including primary score testing, exist. Though their reliability has been questioned, they have content validity and can serve placement purposes. Teacher, peer, and self-evaluation techniques can be applied to holistic approaches. At the other end of the holistic/atomistic continuum are the widely used norm-referenced tests of editing skills, such as the General Education Development Writing Skills Test, Cooperative English Test, Missouri College English Test, McGraw-Hill Basic Skills System, Scott-Foreman Test, Mills' Writing Pre-Test, and the Test of Standard Written English. Critics of these tests cite immeasurable factors that lower the correlation of the quality of a whole essay with the quality of its parts. Recently developed procedures focus on the intellectual processes involved in writing, by measuring syntactic complexity, analyzing linguistic structures, and analyzing the degree of focus. Other recent developments exploit the capabilities of the computer in tabulating specific items such as mature word choice, natural language, and fluency. In addition to describing a pilot study of seven measures of writing ability, this paper describes and evaluates each of the methods of measurement discussed. (AYC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED170014

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Carolyn G. Hartnett

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

MEASURING WRITING SKILLS

Carolyn G. Hartnett

College of the Mainland

Texas City, Texas

Done with assistance from an AIDP Grant

Fall, 1978

JC 790 327

copyright by Carolyn G. Hartnett

MEASURING WRITING SKILLS

TABLE OF CONTENTS

Part One

Introduction: Assessing the Whole or Its Parts	1
Holistic Scales	2
Primary Trait Scoring	5
Peer Evaluation	10
Standardized Measures of Editorial Skills	12
Subtotaling	21
Syntactic Complexity	24
Intellectual Processes	27
Computer-Aided Descriptions	31
Cohesion	34
A Combination Measure	36

Part Two

A Pilot Study Trial of Seven Measures	37
Appendix:	
Composition Tabulation Sheet	43
Intellectual Process Cues	45
Measurement of Cohesion	47
Table of Results	50
Correlation	51
Table of Estimates of Usefulness of Different Measures	54
Bibliography	55

Measuring Writing Skill

Introduction: Assessing the Whole or Its Parts

Atomistic approaches to writing view a composition as a collection of features. In contrast, holistic approaches assume either that the features are too numerous or too complex to assess separately or that the whole of the discourse has an effect beyond the sum of its features. Because that effect is the communicative purpose of the writing, it is easy to argue that a human reaction is absolutely required for any valid judgment of writing quality. If the purpose of writing is to affect a reader's mind, there is no way to judge the writing without a reader's mind. Supplying this second mind for the writer to influence is expensive and variable, however. The reliability of graders has been recognized as a problem since the 1880's.[1] Hirsch contrasts the evaluation modes of Plato and Aristotle; he calls assessment "the single most important snag to practical progress in composition teaching and research."[2] Thus it is worthwhile to consider both holistic and atomistic methods of assessing the quality of writing.

[1]John C. Follman and James A. Anderson, "An Investigation of the Reliability of Five Procedures for Grading English Themes," RTE 1 (1966), 190-200.

[2]E. D. Hirsch, Jr., The Philosophy of Composition, (Chicago: University of Chicago Press, 1977), p. 176.

Holistic Scales

Holistic evaluation of writing includes any method that sorts or rank orders compositions without enumerating their linguistic, rhetorical, or information features.[3] One type of holistic evaluation matches the essays in question with a series of similar model compositions arranged according to quality. Analysis of the scale compositions can show that the graders who arranged the scale had certain criteria in mind. For example, the London scale criteria for imaginative writing were realization of the writer, comprehension of the audience, organization, density of information, and control of written language; and in the Sager Writing Scales, researchers, teachers, or junior high students assign zero to three points for each of these four aspects: vocabulary, elaboration, organization, and structure.[4] Obviously, teachers can develop separate scales for different types of assignments. A dichotomous scale lists features to check as present or absent in the composition. An analytic scale allows a wider range of responses regarding each feature. It could

[3]Charles R. Cooper, "Holistic Evaluation of Writing" in Evaluating Writing: Describing, Measuring, Judging, ed. Charles R. Cooper and Lee Odell (Urbana: NCTE, 1977).

[4]William T. Fagan, Charles R. Cooper, and Julie M. Jenson, Measures for Research and Evaluation in the English Language Arts, (Urbana: ERIC and NCTE, 1975), p. 203. See also London Association for the Teaching of English, Assessing Compositions: A Discussion Pamphlet, (London: Blackie, 1965).

include elaborate descriptions of each feature and definitions of high, middle, and low scores. Although raters doing general impression marking have no such list of specifics, their experience discussing sample papers in detail gives them a mental rubric to account for their reliability.

Other types of holistic evaluation include center of gravity responses, which provide informal feedback to student compositions by summarizing them. This system helps students with revision but is not intended for scoring.[5] Feature analysis and primary trait scoring depend on narrower assessments of features relevant to specific assignments.

Although holistic evaluations are notoriously unreliable, raters with similar backgrounds can be trained to achieve high reliability.[6] The training takes cooperation and time and thus adds to the expense of holistic evaluations, although Godshalk and Diederich assert that raters spending less than two minutes per paper can achieve reliability as high as .90 using some of the holistic scoring guides they describe.[7]

[5]Peter Elbow, Writing Without Teachers, (New York: Oxford Univ. Press, 1973), p. 86.

[6]Paul B. Diederich, Measuring Growth in English, (Urbana: NCTE, 1974). Diederich established a scale by factor analysis.

[7]Fred I. Godshalk, Frances Swineford, and William E. Coffman, The Measurement of Writing Ability, (New York: Educational Testing Service, CEEB, 1966). See also A.E. Myers, Carolyn McConville, and W.E. Coffman, "Simplex Structure in the Grading of Essay Tests," Educational and Psychological Measurement, 26 (1966), 41-54.

Mechanical correctness is usually part of the scoring guide and prompt used in Los Angeles County Schools. Their holistic general impression marking assesses writing competence on a district-wide level. The prompts used for competency testing must meet a list of requirements: (1) be about real life situations, (2) not require students to demonstrate knowledge outside of general experience common to the age and grade level of the test-takers, (3) generate a response that can be completed in twenty minutes, (4) be capable of being completed on an 8½ x 11" sheet of paper, (5) allow for some creativity (for better responses), (6) require specific information, (7) be non-sexist in name, nature, and content, and (8) be of interest to teachers, who are the graders. Los Angeles can get sufficient inter-rater reliability by training teachers for three hours to internalize a scoring guide based on the prompt.

Los Angeles teachers have found they can teach many highschool classes to design a prompt and use it for peer grading. Preparing the prompt helps the students to learn the requirements of good writing; peer grading allows more frequent writing assignments and builds trust. If the teachers were of a mind to, they could incorporate ideas from Judine's collection or even duplicate the ETS Composition Scales by including in their prompt and scoring guide the results of Diederich's factor analysis: ideas and organization doubleweighted, wording, flavor, usage, punctuation, spelling, and handwriting. The major difference between holistic scales and primary trait scoring is that the latter might at times be more narrow.[8]

[8]Division of Curriculum and Instructional Services (Julia M. Gottesman and Marian S. Schilling), A Common Ground for Assessing Competence

College of the Mainland uses a descendent of the Cleveland Rating Scale in its English 131 classes.[9] A copy of it appears in the appendix. It includes percentage weights for organization, development, and coherence, and then subtracts points weighted for mechanical errors (grammar, punctuation, spelling). However, even experienced instructors need much more than the two minutes of the Educational Testing Service raters for mass grading.

Holistic procedures have obvious content validity. They could serve placement purposes. When they involve rank ordering or specific scores, they can measure growth. Diagnostic purposes would require rating with more care for detail. These elaborations increase the administrative difficulty, however.

Primary Trait Scoring

Although most holistic methods of assessment assume that the features of good writing are universal, primary trait scoring grows from a theoretical base that sees variation beyond the universal qualities.

---

in Written Expression, (Downey, CA: Los Angeles County Supertinendent of Schools, 1978). See also Paul B. Diederich, "How to Measure Growth in Writing Ability," English Journal 55, (April, 1966), 435-49; Sister M. Judine, I.H.M., A Guide for Evaluating Student Composition, (Urbana, Ill., NCTE, 1965).

[9]Follman, op. cit., p. 193.

Considerable evidence exists that the mode of a composition (such as description, narration, exposition, or persuasion) affects the way it is written. The effect extends even to seemingly objective details of syntactic structure.[10]

[10]David P. Boder, "The Adjective-Verb Quotient: A Contribution to the Psychology of Language," Psychological Record 3 (1940), 310-343; Ellen Frogner, "Problems of Sentence Structure in Pupils' Themes," English Journal 22 (1933), 742-749; J . C. Seegers, "Form of Discourse and Sentence Structure," Elementary English Review 10 (March, 1933), 51-54; Gerald Lloyd Kincaid, "Some Factors Affecting Variations in the Quality of Students' Writing," Diss., Michigan State University 1953, as described in Richard Braddock, Richard Lloyd-Jones, and Lowell Schoer, Research in Written Composition (Champaign: NCTE, 1963), pp. 83-95; Lois V. Johnson, "Children's Writing in Three Forms of Composition," Elementary English 44 (1967), 265-269; Donald Richard Bortz, "The Written Language Patterns of Intermediate Grade Children When Writing Compositions in Three Forms: Descriptive, Expository, and Narrative," DAI 30 (1970), 5332 A (Lehigh); D. A. Nietzke, "The Influence of Composition Assignment Upon Grammatical Structure," DAI 32 (1972), 5476 A; Mike Pope, "The Syntax of Fourth Graders' Narrative and Explanatory Speech," RTE 8 (1974), 219-227; Harold Rosen, "An Investigation of the Effects of Differentiated Writing Assignments on the Performance in English Composition of a Selected Group of 15/16 Year Old Pupils,"

Recent research analyzes the relationship of syntactic structure to the sex of the writer.[11] Variables of audience, topic, and tone or style have obvious effects.

The National Assessment of Educational Progress scored a mixture of 1969 and 1974 essays both holistically and descriptively and reported: "The nature of holistic scoring is such that one never knows precisely why a paper received the rating it did; the scorers themselves do not reflect on their decisions, only knowing that a particular paper is better than some but not as good as others. Consequently, this method of scoring provides a very reliable ranking of essays--a ranking most trained readers would endorse--but it tells us nothing about the papers except that some are better than others. Descriptive scoring of essays provides considerable information about essays, but it does not tell anyone how good a particular essay is when read. This study combines both systems in order to characterize specific writing . . . . The declines in holistic scores reveal as much about scorers as they do about students. Language is always changing, and scorers may prefer standards of written expression that are becoming outmoded. New standards are

---

The Development of Writing Abilities (11-18), James Britton, Harold Rosen, et al. (London: Macmillan, 1975), pp. 2-3.

[11]Mary P. Hiatt, "The Feminine Style: Theory and Fact," CCC 29 (October, 1978), 222-226.

certainly different, but they may not be worse in any defensible sense."[12]

Recognizing the differences, primary trait scoring looks at only the characteristics unique to a specific assignment. The primary trait score describes only the specific aspects it was designed to assess. Usually an exercise earns one point if the trait sought is absent, two points for its presence, three for competence, and four for excellence.[13] It requires careful preparation of the assignment and of the scoring guide. The theoretical base of assumed variation limits its extension to other types of writing. Thus it is necessary to take great care in preparing the assignment to make its application as broad as possible. Extensive, careful preparation of a primary trait assignment and scoring guide was justified by its broad use in the National Assessment of Educational Progress. The use of Primary Trait Scoring for placement and diagnosis is valid only where the design relates to such use. For example, if we had a Piagetan developmental sequence of writing tasks,

[12]Writing Mechanics, 1969-1974: A Capsule Description of Changes in Writing Mechanics, Writing Report No. 05-W-01 (Denver: National Assessment of Educational Progress, 1975). See also Write/Rewrite: An Assessment of Revision Skills, Writing Report No. 05-W-04, 1977, and Explanatory and Persuasive Letter Writing, Writing Report No. 05-W-03, 1977, pp. 3, 43.

[13]Ina V. S. Mullis, The Primary Trait System for Scoring Writing Tasks, (Denver: National Assessment of Educational Progress, 1975), p. 6.

success on an assignment on one level might place the student within the sequence. Thus primary trait scoring might also diagnose learning needs within the sequence. Without such an assumed sequence, however, primary trait scoring is limited to measuring its specific aspect within the weight of the entire paper. Two similar tasks could trace growth for an individual teacher or for a national researcher.

Cultural bias that distorts responses becomes more likely as the assignment gains in specificity; yet specificity improves motivation and reliability of scoring.[14] In a single classroom assignment, extended breadth of application is not necessary. A teacher can use primary trait scoring to show students the strengths of their writing. The teacher could easily write specific behavioral objectives stating the primary traits an assignment must show. Preparing a scoring guide provides a detailed and exact description of the performance expected on a specific rhetorical task. Primary trait scoring can ignore features of syntactic fluency, organization, diction, mechanical correctness, etc., to get at questions searching the depth of thought. Analysts have found that mechanics correlate with length and quality of developed ideas.[15]

[14]Richard Lloyd-Jones, "Primary Trait Scoring," Evaluating Writing, ed. Charles R. Cooper et al. (Urbana: NCTE, 1977), p. 42.

[15]Henry B. Slotnik and W. Todd Rogers, "Writing Errors: Implications about Student Writers," RTE 7 (1973), 387-98.

A primary trait scorer needs a firm grasp of the scoring guide to prevent the influence of unmeasured variables from decreasing reliability. Questions of judgment arise when a poor writer approaches the target without hitting it exactly.

Peer Evaluation

Who is capable of performing adequate holistic evaluations? Although the scaling methodologists of Educational Testing Service (ETS) stress the need for training readers with similar backgrounds in order to achieve reliability, they do not state minimum competencies of the graders.[16] In 1976 E. D. Hirsch copyrighted "A Method for Forming Tests to Certify Assessors of Writing Ability" that requires candidates to assign grades to writing samples. The criteria is readability. A system of quizzing audiences on the content of the samples determines the target grades. Candidates able to assign target grades receive certification as assessors for writing done by a specific age group. Certified assessors can then create new tests.[17]

Although certified assessment ability would certainly be useful for a teacher, good students also value this skill. In fact, all students of writing skills need an image of good writing, an image gained from reading. Beaven cites six assumptions that underlie student partici-

[16] Diederich, op. cit., and Godshalk, op. cit.

[17] Hirsch, op. cit.

pation in writing assessments.[18] Behind the assumptions are premises that growth in writing is a highly individualistic process related in many ways to growth in personality and procedures designed to foster such growth. One assumption is that growth in writing occurs slowly, which means changes are more noticeable at two or four year intervals than during the course of a semester or quarter.[19] Growth does not occur in isolation. Moffett believes that maturity alone will suffice for developing the cognitive structures of mature writing.[20] Students need sensitization to themselves and their surroundings, awareness of their audience, and techniques for invention and revision. Maturation of writing skill takes goal-setting and risk-taking. Students need a climate of trust.

These assumptions lead to individualized goal setting, which is especially useful at the beginning of the year within the single classroom. Self-evaluation grows from it. Students can use checklists or any of the scales for holistic rating. They can evaluate the length of

[18]Mary H. Beaven, "Individualized Goal Setting, Self-Evaluation, and Peer Evaluation," Evaluating Writing, ed. Charles R. Cooper et al. (Urbana, NCTE, 1977), pp. 135-156.

[19]Walter Loban, Language Development: Kindergarten through Grade Twelve, Research Report No. 18, (Urbana, NCTE: 1976).

[20]James. A. Moffett, Teaching the Universe of Discourse (Boston: Houghton-Mifflin, 1968), pp. 76-83, 163.

time they expended, their improvements, their strengths (which they may overlook otherwise), and their weaknesses (which show that imperfection is acceptable and a source for learning).

Peer evaluation has many of the advantages of self-evaluation and avoids the problem of excessive dependence on the teacher as a single audience.[21] It realistically applies writing to test its communicative function. Sager found that sixth graders could use a scale she developed to improve their writing as much as when only teachers evaluated student work.[22]

Despite its values, peer evaluation presents problems. Many teachers themselves doubt its face validity, objectivity, and reliability. Administration is difficult: It takes extensive class time and usually a teacher trained in group dynamics. Although it could certainly offer useful hints on placement, diagnosis, and growth, it would be somewhat awkward. Its value for these purposes has not been tested and proven.

Standardized Measurements of Editorial Skills

Holistic evaluations are at the opposite end of the continuum from the widely-used norm-referenced tests of editing skills. An example of an atomistic test used only to certify competency is the "GED Writing Skills Test." The examinees need to recognize errors and make choices

[21] James Britton et al., The Development of Writing Abilities (11-8) (London: Macmillan, 1975). See also Beaven, op. cit.

[22] Sager, op. cit.

of effectiveness. Skills include spelling, punctuation and capitalization, grammar and usage, diction and style, sentence structure, and logic and organization.[23]

The Cooperative English Test has broader uses: prediction, placement, assessment, and evaluation. It offers multiple choice questions on spelling, diction (such as double negatives), subject-verb agreement, comma splices, case, and use of semicolons and commas. On three-quarters of the questions, students are to detect errors.[24] The Cooperative Sequential Tests of Educational Progress (STEP) in Writing has students find errors or best choices regarding spelling, usage or diction, effectiveness, logic, punctuation, etc.[25]

The Missouri College English Test is intended for both placement and achievement measure, although its authors recognize it can measure only a sample of "those aspects of learning commonly accepted as being important and measurable."[26] In sixty items students hunt for errors of

[23]GED Testing Program, Information Pamphlet on the New GED Tests, (Washington: American Council on Education, 1975).

[24]Cooperative Test Division, Cooperative English Tests: Technical Report, (Princeton: Educational Testing Service, 1960).

[25]Cooperative Test Division, Cooperative Sequential Tests of Educational Progress, (Princeton: Educational Testing Service, 1957).

[26]Robert Callis and Willoughby Johnson, Missouri College English Test Manual, (New York: Harcourt, 1965), p. 14.

punctuation, capitalization, grammar, and spelling. In thirty other items students choose the best of four sentences or arrange given sentences in a paragraph.

The McGraw-Hill Basic Skills System writing test measures "skill in written communication" mainly for purposes of placement in college English courses.[27] Students find errors in capitalization, punctuation, and grammar; identify sentence types and fragments; and choose topic, developing, and concluding sentences appropriate for given paragraphs. The publishers did not correlate scores with any criterion except their own texts. CTB/McGraw-Hill published also English test booklets with separate sections on sounds, comprehension, mechanics, and organization of ideas. In the last, students choose items from lists to construct outlines of the structure of a composition.[28]

In 1978 Scott, Foresman published a multiple-choice diagnostic test to accompany its Handbook of Current English and Workbook of Current English. The Handbook states, "The Diagnostic Test is intended to test students' mastery of grammatical and rhetorical skills."[29] In fifty

[27]Alton L. Raygor, ed., McGraw-Hill Basic Skills System Writing Manual, (Monterey, Calif: CTB/McGraw-Hill, 1970).

[28]Educational Skills Tests, College Edition, (Monterey, Calif: CTB/McGraw-Hill, 1971).

[29]Jim W. Corder, Handbook of Current English, (Glenview, Ill.: Scott, Foresman, 1978), p. xi.

items, students select one correct sentence from four options; three of the choices have errors in grammar, punctuation, or spelling. In the thirty-five items of the rhetoric section, students must choose the best topic, wording, sentence, or paragraph. Best means most effective or most appropriate for a 500-word theme.[30] Although this test can be machine scored, its answers do not satisfy some English instructors any better than do the other editorial tests, a sample of which are listed here. Liberal and conservative approaches result in differences here and also in marking student writing directly.[31]

Mills has reported that her proofreading test called "Writing Pretest" is not as good a predictor of writing style as her "Sentence Pretest," which had students follow examples analyzing pairs of sentences and generating similar ones.[32] She developed these tests for placement, diagnosis, and measurement of growth in the text, Commanding Paragraphs,

[30]Hugh H. Paschal, Diagnostic Test to Accompany Handbook of Current English and Workbook of Current English, (Glenview, Ill.: Scott, Foresman, 1978).

[31]Richard Braddock, Research in Written Composition, (Champaign: NCTE, 1963), p. 16.

[32]Helen Mills, "Language and Composition: Three Mastery Learning Courses in One Classroom," Journal of Basic Writing, Fall, 1976, pp. 44 - 59.

which requires tenth grade reading ability. Although a trained grader spends only five minutes on the "Sentence Pretest," Mills continues to use the "Writing Pretest" for its administrative ease.

The College Entrance Examination Board "Sentences Test" consists of forty sentences. Each has a "no error" option and four spots of possible error in grammar, usage, word choice, idiom, capitalization, and punctuation.[33]

A more recent atomistic test of the editorial type is the Test of Standard Written English (TSWE), which Educational Testing Services published in 1974 to supplement the Scholastic Aptitude Test (SAT). Its fifty questions have students either "recognize writing that does not follow the conventions of standard written English" or "choose the best way of phrasing a sentence."[34] It tests grammar (such as subject-verb agreement), usage, word choice, idiom, comparison, coordination, and subordination, but it does not check on spelling, capitalization, formal grammatical terminology, or much punctuation. Some professionals have objected to this content of the TSWE.[35]

[33]College Entrance Examination Board, Comparative Guidance and Placement Program: Sentences, (Princeton: ETS, 1967).

[34]Educational Testing Service, The Test of Standard Written English: A Preliminary Report, (Princeton: ETS, 1975).

[35]Sandra Clark, "Problems with the Test of Standard Written English," CLAC - 4, (March, 1978), 10-13.

The publishers of TSWE show that it correlates with grades in college English courses and with writing performance.[36] Their purpose for TSWE is placement only, not to certify achievement. Breland suggests that individual instructors who want to teach the kinds of things tested in TSWE should prepare their own posttests of progress. He warns that teaching to raise TSWE scores is difficult; five points per semester is the maximum that can be expected. Teaching to raise the rating of an essay is even more difficult. In this respect, these measures may resemble intelligence tests.

Breland reports no significant differences in correlation for sex or minority status when all minorities are combined, but he recommends further research on separate minorities. For minority students and low-ability women, TSWE predictions tend to be higher than actual performances; in addition, TSWE underestimates the actual performance of high ability women.[37]

[36]Hunter M. Breland, A Study of College English Placement and the TSWE, Project report 77-1 (Princeton: ETS, 1976).

[37]Hunter M. Breland, Group Comparison of the Test of Standard Written English, Research Bulletin RB-77-15, (Princeton: ETS, 1977).

At Yale, students who scored low on TSWE wrote poor essays, but high scorers did not always write well.[38] TSWE correlated best with subsentence and sentence parts of an English composition rating. It correlated next best with SAT verbal scores. The study concluded that TSWE was adequate and efficient for placement and more economical than an essay. It also revealed that justifiable objections to the quality of Yale student writing were related to dullness and the lack of logic or thought rather than to mechanical requirements.

Some users of TSWE see in it cultural bias and a discriminatory hidden agenda that leads to overemphasis of conventional linguistic etiquette.[39]

The Conference on College Composition and Communication (CCCC) has passed several resolutions on testing; one in 1974 stated: "Resolved, first, that CCCC protest the inclusion of an objective usage test in the Scholastic Aptitude Test, on the grounds that such tests are a measure of copyreading skill rather than a measure of student

[38]Judith D. Hackman and Paula Johnson, Yale College Freshmen: How Well Do They Write? (New Haven: Yale Univ., 1976); see also by the same authors, "Yale: How Well do Freshmen Write? Implications for Placement and Pedagogy," College and University, 53 (Fall, 1977) 81-99.

[39]Keturah Funkhouser, "TSWE: 'T is Not for Me," CLAC-3, (Nov., 1977), 13-16.

ability to use language effectively in connected discourse of their own composing; such tests place emphasis on mechanical matters of spelling, punctuation, and conventions of usage, rather than on clarity, appropriateness, and coherence of thought; such tests tend to discriminate against minority students whose linguistic experiences often lead them to choose answers different from those expected by the test-makers; and the inclusion of such a test may encourage secondary English teachers to teach toward the test at the expense of matters more fundamental to effective writing and sophisticated reading; and second, that CCCC encourage its members to resist the use of usage scores in the admission and placement of students."[40]

In 1978, CCCC resolved that: "No student shall be given credit for a writing course, placed in a remedial writing course, exempted from a required writing course, or certified for competency without submitting a piece of written discourse." The resolution requires consideration of the complexities of both multiple choice and essay tests and further study of the entire issue of testing.[41]

[40]"Resolution No. 6" and "Resolution No. 7," CCC 25, (Oct., 1974), 339.

[41]"Resolution No. 1," CCC, (Oct., 1978), 309.

Braddock calls it charitable but inaccurate to say these objective tests measure editing ability; they ask for proofreading skills.[42] They do correlate with success in courses well enough for screening when teachers can transfer students after reading their writing; but what they evaluate is not actual writing ability. Fifteen percent of the students with TSWE scores above the 95th percentile write essays that are below average.[43] Test questions deal with the facets of English that are easiest to test.[44] Yet even here, problems arise and professionals disagree. Hooks points out obvious errors, weaknesses, and a few strengths of objective tests.[45]

Tests tell students, teachers, and the public what the tester believes is worth testing. If the content is patently trivial, the whole course is trivialized, demeaned, belittled, dehumanized. However, the College Board recognized its leadership role and issued guides for teachers

[42]Richard Braddock, "Evaluation of Writing Tests," Reviews of Selected Published Tests in English, ed. Alfred H. Grommon, Richard Braddock, J. N. Hook, William A. Jenkins, Walter Loban, and Alan C. Purves, NCTE Committee to Review Standardized Tests, (Urbana: NCTE, 1976).

[43]Educational Testing Service, Guide to the Admissions Testing Program 1978-79, (Princeton: CEEB, 1978), pp.9, 16.

[44]Task Force on Measurement and Evaluation in the Study of English, Alan Purves, Chairm Common Sense and Testing in English, (Urbana: NCTE, 1975).

[45]J. N. Hooks, "Tests on the English Language," Reviews of Selected Published Tests in English, ed. Alfred H. Grommon, (Urbana: NCTE, 1976), pp. 76-117.

planning advanced placement courses, directly influencing course content. The actual test is two-thirds essay: "the examination is unique because it requires writing as a direct measure of the students' ability to read and interpret literature and to use other forms of discourse effectively."[46]

Subtotaling

If it seems strange to call a direct writing test unique, the complexities that justify it need explanation. There are logical reasons why evaluating the whole is more difficult than evaluating parts and summing them. These complicating factors lower the correlation of the quality of a whole essay with measurements of the quality of its parts, such as these mechanical tests just described and devices that will be described in the following sections. The biggest problem is that what we need to measure is largely what Polanyi called "tacit knowledge," the unmeasureables that underlie competence.

Researchers have pointed at many complicating factors. Although correlation is not causation, in Loban's study proficiency in language accompanied reasonable affluence.[47] Ajay found that adding variables

[46]Educational Testing Service, Advanced Placement Course Description: English Composition and Literature, (Princeton: CEEB, 1978), p. 3.

[47]Loban, op. cit.

from content improves the prediction of writing quality from style alone.[48] Braddock cites claims from typewriter manufacturers that typists spell better and write more than other students, but some research on this point is mixed.[49]

Sanders concluded that timed impromptu conditions and assigned topics and modes limit both motivation and quality.[50] Woodfin found that a longer time allowance improves quality for third graders.[51] In a 1963 description of the standard research models, Braddock warned, "Even if the investigator is primarily interested in nothing but grammar and mechanics, he should afford time [for writers to plan central idea, organize, and support as well as edit and proofread]. Otherwise their sentence structure and mechanics will be produced under artificial circumstances."[52] Furthermore, Diederich, Kincaid, and others conclude

[48] H. B. Ajay, Strategies for Content Analyses of Essays by Computer, Dissertation Abstracts International 34 (1973), 2375A.

[49] Braddock, op. cit., p. 51.

[50] Sara E. Sanders and John H. Littlefield, "Perhaps Test Essays Can Reflect Significant Improvement in Freshman Composition: Report on a Successful Attempt," RTE 9, (1975), 145. Footnote number 10 cites other evidence on the influence of mode.

[51] M. J. Woodfin, "The Quality of Written Expression of Third Grade Children under Different Time Limits," Journal of Experimental Education 37, No. 3, (1969), 89-91.

[52] Braddock, Research, p. 9.

that at least two writing samples are needed to allow students a chance to do their best.[53]

When we do bother to collect adequate writing samples, and when we are able to take the time, we need valid, reliable, objective ways of assessing important features. We need methods that satisfy Emig's complaint, "Most of the criteria by which students' school-sponsored writing is evaluated concerns the accidents rather than the essences of discourse--that is spelling, punctuation, penmanship, and lengths rather than thematic development, rhetorical and syntactic sophistication, and fulfillment of intent."[54] She found as she investigated how students actually wrote (in contrast to their teaching) that, for example, construction of a formal outline did not correlate with highest quality writing. Analyses of actual practices such as Emig and Christiansen have done lead to changes in the content of composition courses.[55] Humanities as well as physics follows Heisenberg's principle: measurement influences and changes what is being measured.

Mathematical methodology exists to count the amount of information in each word and sentence.[56] It requires determining previously

[53]Kincaid, op. cit., Diederich, 1974, p. 34.

[54]Janet Emig, The Composing Processes of Twelfth Graders, (Urbana: NCTE, 1971), p. 93.

[55]Francis Christensen, Notes Toward a New Rhetoric, (New York: Harper and Row, 1967).

[56]Klaus Weltner, The Measurement of Verbal Information in Psychology and Education, tr. Barbara M. Crook (New York: Springer-Verlag, 1973).

known information in order to label new data. It must recognize that the English language is more than a Markov series of left-to-right probabilities and that language has purposes beyond conveying information. Eventually this process may have classroom use, but for now it concerns only researchers and planners of textbooks.

Syntactic Complexity

A method of measuring syntactic complexity of sentences has value for the researcher as well as for the classroom teacher who wants to determine readability or assess student writing. Several methods exist. Dupuis, for example, presented a list of twenty-seven possible transformations: addition, deletion, reordering, combining, and variations.[57] Her inter-rater reliability was .94. Endicott proposed a scale of syntactic complexity based on the proportion of units, called co-memes, in the surface structure of a sentence. His scale begins with a sentence lacking all the optional transformations except tense morphemes.[58] Another test has students expand kernals of a synopsis of The Adventures of Huckleberry Finn. Analysis of their responses reports the subordination ratio, words per subordinate clause and T-unit, and mean depth of embedding.[59] This method and

[57]Mary M. Dupuis, "Transformational Analysis of Compositions (TAC)," Measures, ed. Fagan, pp. 193-94.

[58]Anthony L. Endicott, "A Proposed Scale for Syntactic Complexity," RTE 7, (1973), 5-12.

[59]Fritz Dauterman, "Syntactic Maturity Test for Narrative Writing," described in Fagan, pp. 188-89.

other ways of analyzing syntactic complexity are based on the work of Kellogg Hunt and Francis Christensen. Hunt showed that the length of the mean minimal terminable unit of a sentence (T-unit) increases with maturity.[60] He thus provided a tool for describing syntactic fluency that has led to techniques of teaching that use sentence-combining activities. Another researcher, Dixon, compiled a list of indexes or predictors of syntactic maturity: mean T-unit length, words in final free modifier position, instances of final free modifiers, words in intra-T-unit coordinations, instances of such coordination, and total words and instances of free modifiers in all positions.[61] He found Hunt's T-unit the best single index, accounting for 44% of the distinction in student writings. Free modifiers in all positions accounted for 30%; and a combination of T-unit and free modifiers can justify 58%. Much can be done with this type of analysis. For example, Wolk had his students explore Christensen's hypothesis of the importance of final free modifiers by comparing their own essays with twelve professional information articles.[62]

[60]Kellogg W. Hunt, "Early Blooming and Late Blooming Syntactic Structures," Evaluating Writing, ed. Cooper, pp. 91-106.

[61]Edward Dixon, "Syntactic Indexes and Student Writing Performance," Elementary English 49, (May, 1972), 714-716.

[62]Anthony Wolk, "The Relative Importance of the Final Free Modifier: A Quantitative Analysis," RTE 4, (1970), 59ff.

Measures of syntactic complexity demonstrate growth over long periods of time, several years. The question of the value of this complexity has a classic answer from Moffett, "Children's sentences must grow rank before they can be trimmed."[63] Moffett also shows that vocabulary development makes a single mature word substitute for a complex clause, and thus a non-superficial theory of complexity must consider semantics.

Some research fails to show that longer T-units correlate with good writing.[64]

Administration of measures of syntactic complexity are time consuming in actual counting, in obtaining an adequate sample, and in application. A warning was mentioned earlier that the mode of discourse affects the syntax and thus the length of T-units.[65] One might hesitate with placement or diagnosis based only on T-unit length because of factors like context and effect, which influence syntax; yet there is no doubt that syntactic analysis provides a base for much good classroom instruction.

[63]James B. Moffett, Teaching the Universe of Discourse, (New York: Houghton Mifflin Company, 1968), p. 172.

[64]Robert R. Potter, "Sentence Structure and Prose Quality," RTE 1 (Spring, 1967), 17ff. See also R. G. Martin, "A Prediction Formula for a Sample of 'Good' Writing," Dissertation Abstracts International 29, (1968), 1221A. Other studies also exist. See entire issue RTE 12 (October, 1978), 3.

[65]See footnote number 10.

Intellectual Processes

Maturation develops a child's ability to abstract hierarchically. Moffett sees "a parallel between qualifying thought and elaborating sentence structures" as they develop together.[66] Cognitive development over the years is more necessary than grammatical knowledge in stimulating syntactic growth and elaboration, according to Loban after a longitudinal study.[67]

One of the early prophets of sentence-combining as a method for enhancing syntactic fluency was John Mellon, who looked at the prerequisite skills for sentence combining. He called syntactic maturity a statistical artifact and related increasing elaboration to conceptual development.[68] If students lack cognitive development, he said, sentence combining adds only fog. Dominant noun phrases cannot be more complex than the concept in the writer's mind. As the concept grows, so does syntactic memory. Organizing, chunking, abstracting processes depend on prior skills of memory and attention, which develop through maturation.

[67] Loban, op. cit., pp. 35-36. [66] Moffett, op. cit., p. 77.

[68] John C. Mellon, "Issues in the Theory and Practice of Sentence Combining: A Twenty-Year Perspective," in Sentence Combining and the Teaching of Writing, ed. Donald Daiker, Andrew Kerek, Max Morenberg (Akron: Univ. of Akron Dept. of English, forthcoming). See also Mellon's earlier work, Transformational Sentence-Combining: A Method for Enhancing the Development of Sentactic Fluency in English Composition, (Champaign: NCTE, 1969).

Piaget characterized the concrete operational student's thinking as capable of decentration, of dynamic transformation, and of empirical reverses.[69] These abilities are exactly what sentence combining requires: looking objectively at kernal data to re-arrange it effectively. Children's speech develops from social purposes through egocentric meanings to inner goals, according to Vygotsky.[70] These stages resemble but do not exactly parallel Kinneavy's aims for discourse: reference or explanation (subject-oriented), persuasion (oriented to the audience), and self-expression and literature (oriented to code or media).[71] However, Moffett showed how the modes of discourse develop in normal maturation.[72] Another good explanation of the process is Britton's.[73]

Perhaps we need to look at intellectual processes in order to answer Ney's question: How do sentence combining exercises affect the

[69]Herbert Ginsburg and Sylvia Opper, Piaget's Theory of Intellectual Development (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969).

[70]Lev Semenovich Vygotsky, Thought and Language, tr. and ed. E. Hanfmann and G. Vakar, (1934, Cambridge: MIT Press, 1962).

[71]James Kinneavy, A Theory of Discourse: The Aims of Discourse, (Englewood Cliffs: Prentice-Hall, Inc., 1971).

[72]op. cit.

[73]James Britton, Language and Learning (Middlesex, England: Penguin Books Ltd., 1970), pp. 203-237.

psycholinguistic abilities of students so that observable changes appear in writing?[74] How does quality of writing relate to intellectual maturation, which develops through gradual increments, and how does it relate to intellectual capacity, which traditionally changes very little? If it is as hard to raise a TSWE score as an IQ, it may be more to the point to teach and test intellectual processes directly, rather than indirectly.[75] The Yale study mentioned above found dullness and lack of logical thought a more serious problem than mechanical correctness.[76] The English education theoretician Kinneavy has heretically questioned the competence of English teachers to teach and evaluate logic.[77]

Aristotle taught invention; today we teach pre-writing. One approach to invention is Pike's tagmemic heuristic procedure. Pike recognizes the relationships in the hierarchy of linguistic elements. He examines anything, linguistic or non-linguistic, as particle, wave, and

[74]James Ney, "Notes Toward a Psycholinguistic Model of the Writing Process," RTE (1974), 159-169.

[75]Breland, A Study, pp. 50-51

[76]Hackman, op. cit.

[77]James Kinneavy, "Sentence Combining in a Comprehensive Language Framework," in Sentence Combining and the Teaching of Writing, ed. Donald Daiker, Andrew Kerek, and Max Morenberg, (Akron: Univ. of Akron, forthcoming).

field, all with contrasts, variations, and distributions--a ninefold examination.[78] Although Pike developed some of the important technical procedures to analyze linguistic structures, he also stressed, "Beyond the linguist lies the artist."[79] An ideal test would measure both linguistics and artistry.

The most specific published method of analyzing intellectual processes in student writing is Odell's.[80] Odell develops focus, the first of the six processes, from Pike. Grammatical subjects reflect focus, the heirarchical level of abstraction. Second, connectors, comparative and superlative forms, negatives, negative affixes, and lexical items are linguistic cues to the use of contrast. The cues to classification and pattern recognition are predicate nominatives, labels of examples and instances, and lexical terms of resemblance or class. Synonyms for change, become, begin, and stop all indicate change, the fourth process. For placement in physical context the linguistic cues are nouns referring to geographical locations, objects in physical settings, or sensory properties of physical settings. To show time

[78]Richard E. Young, Alton L. Becker, and Kenneth L. Pike, Rhetoric: Discovery and Change, (New York: Harcourt, Brace & World, Inc., 1970), p. 127.

[79]Kenneth L. Pike, "Beyond the Sentence," CCC15, (1964), 129-135.

[80]Lee Odell, "Measuring Changes in Intellectual Processes," Evaluating Writing, ed. Charles Cooper and Lee Odell, (Urbana: NCTE, 1977), pp. 105-132.

sequence, English has adverbial elements noting existence before, during, or after a stated moment. Logical sequences use if...then constructions and words implying a cause-effect relation, such as because, therefore, since, and consequently.

Odell recognizes that although counting individual cues is inaccurate, they direct attention to the processes that need tabulation. His independent scorers achieved 88% agreement. Yet he wonders about a need for more subcategories and finer discriminations. He asks about a possible sequence of different subcategories or types of intellectual processes that develop at different ages. Such a sequence would parallel the syntactic development that Hunt found, exemplified by the temporary flourishing of coordination. Odell suggests that identifying intellectual processes can distinguish quality of writing, diagnose problems, and measure growth. He considers intellectual analyses as complementary to analyses of syntax and of semantic choices.

Computer-Aided Descriptions

Computers can assist assessment by tabulating a variety of specific items. One area deals with appropriate mature word choice by comparing the writer's choices with standard frequency indexes. A system by Finn requires typing student themes on IBM cards and programming the computer to print alphabetical lists of all the different words used in each theme

and their frequencies.[81] Finn points out that complicating factors make necessary additional lists of topic-imposed words and of words that are rare in writing but not mature (proper nouns, contractions, and slang, for example). Some common words appear equally at different grade levels and do not distinguish writers; they need no special attention. The list of remaining unclassified word choices may bring in irrelevant ideas. Finn suggests that teachers could use that list to analyze the coherence of the writing.

One of the well-developed computer-assisted instructional programs is JOURNALISM. In it Bishop analyzes natural language.[82] He relates the computer to specific exercises by supplying it with an instructor's list of key words, arranged in their best order. JOURNALISM does stylistic analysis by reporting variety in sentence length and structure overuse of articles, passives, adjectives, and adverbs. It checks spelling by printing any word not on a list of 17,000 words or a special topical supplement. In addition, JOURNALISM performs housekeeping duties by recording student progress.

The work of Christiansen, Hunt, and Mellon underlies WORDS and COUNT, by Webb, programs that resemble Bishop's JOURNALISM. Webb's COUNT reports on the number of words, sentence and paragraph lengths,

[81]Patrick J. Finn, "Computer-Aided Description of Mature Word Choices in Writing," Evaluating Writing, ed. Cooper (Urbana: NCTE, 1977), pp. 69-90.

[82]Robert L. Bishop, "Computing in the Teaching of Journalistic Skills," On-Line 3, (May, 1974), no. 3, 5-12.

ranges, and deviations. WORDS compares the occurrence of thirty key sentence unit connectors and the occurrence of initial, medial, and final free modifiers with their frequencies in fifty random articles published in magazines in 1972.[83]

Another approach is Slotnik's. Slotnik distinguished the measures a computer can make, called proxes, from what he calls trins, the intrinsic qualities that interest human judges. For example, a computer can tabulate the proxes of total number of words, the different words, the commas, and the function words. These proxes are aspects of fluency; and fluency relates to the trin of quality of ideas and generalization. Mean and standard deviation of word size, head gerunds and past participles, and absence of uncommon misspellings relate to diction in this plan. Colons, semicolons, quotation marks, and parentheses are uncommon punctuation marks related to the trin of complex thought and emphasis. Other trins measured are spelling, sentence structure, and paragraphing.[84] The validity of the proxes needs proving.

A student with a printout from any of these computer programs would have raw material to plan revisions. Printouts can suggest to

[83]E. Jackson Webb, Computer Assisted Instruction Program in Writing, (Olympia, Wash: Applied Linguistics, The Evergreen State College, 1973).

[84]Henry B. Slotnik, "Toward a Theory of Computer Essay-Grading," Journal of Educational Measurements 9, (1972), 253-63.

teachers the variables needing attention. Computer tabulations could also provide a backup to a human evaluator, like a second reader. They have the reliability and the culture-fair quality of their program. To measure growth in the specific aspects they count, they are ideal; and an interpreter could use them to diagnose some problems. Placement by computer is a long way off, even for specific aspects. Finn has not yet normed even his vocabulary counts, and every topic assigned would need separate work. The content validity of the computer programs depends on their theoretical assumptions and on their uses. The biggest disadvantage now is programming cost. Extensive use, however, might justify it.

Cohesion

Coherence is a quality desired in all writing. It relates to both intellectual processes and to fluency with language, but it does not directly indicate logic and depth of thought, syntactic complexity, or correctness. Thus it is not complete as a measure of writing quality, but, probably like intelligence, it correlates. To count the cohesive ties in a piece of writing, I adapted a classification system developed by Halliday in 1973. Halliday defined a cohesive tie as a signal sending the reader to another sentence for its meaning. The test of a cohesive tie is finding in the other sentence the word or words to which it refers. As used here, the term cohesion excludes words that relate the text to the real world; it also excludes the structural relations within sentences

that give each well-formed sentence its own inner coherence.[85]

Fluent writers will be aware of their many options and be able to choose a variety of them, but perhaps avoid too much variety. The fifteen types of cohesion include three types of reference: third person personal pronouns, demonstratives (including the), and comparatives. Substitution and ellipsis can relate to a noun (by ones, same, many, mine, more, etc.), to a verb (with do or by omitting tense, mood, or voice markers), or to a clause (with so or as an answer to a question). Conjunctions can add, contrast, continue, show cause, or relate to time and summation. The four lexical types of cohesion are repetitions of the same word; synonyms, hyponyms, or related words; higher categories; and general nouns. (A more complete set of definitions and examples appears in the appendix.) It would be rare to write coherently without at least lexical ties.

Counting cohesive ties resembles counting intellectual processes with Odell's linguistic cues. Cohesion, however, may be more related to language fluency that to only intellectual process or depth of thought. In both counts, the range of types seems more important than the total of actual tokens. Within a short piece of writing (400 words), cohesion seems to have a ceiling of about seven to nine of the fifteen different

[85]Michael A. K. Halliday and Ruqaiya Hasan, Cohesion in English, (London: Longman, 1976).

types. Poor writers use few of the options, but good writers do not use too much variety.

Within the limits stated, a count of cohesive ties has some face validity, and it can measure growth. Most students do use more cohesive ties at the end of a remedial writing course than they do at the beginning. A trained reader could count the cohesive ties quickly, but not faster than a two minute general impression marking. While cohesive ties are objective linguistic cues and may correlate highly with writing quality, they are no more than an index of it. The greatest value of a count of cohesive ties might be to supplement another incomplete measure.

A Combination Measure

Many schools use both a writing sample and an objective test on mechanical and grammatical features. Cooper's definition of holistic marking excludes procedures that enumerate linguistic, rhetorical, or informational features.[86] Thus we must consider separately Myklebust's Picture Story Language Test. It has a purpose different from the tests described above. It diagnoses disorders of written language at a clinical level, although it has been used for placement in remedial college courses. It consists of a stimulus picture for the student to write a story about, a syntax scale, a productivity scale, and an abstract-concrete scale. The syntax score records addition, omission, substitution, and word order

[86] Cooper, op. cit., p. 4.

errors in punctuation, word usage, and word endings. It ignores spelling.[87] Myklebust reported, "it was our constant feeling that the A-C [Abstract-Concrete] Scale would be the most unreliable. To our amazement, this is not at all the way it worked out statistically--the most troublesome scale was that of Syntax."[88] Scoring Myklebust's test takes at least fifteen minutes of secretarial time. (Instructors quibble more over alternate intentions for the syntax score.) Since alternate forms do not exist, it cannot be used to measure growth. Although norms have been collected throughout the world for children, they do not extend into adulthood despite rapid changes expected between ages seventeen and twenty-one.

A Pilot Study Trial of Seven Measures

Appraisal testing at College of the Mainland includes the Nelson-Denny reading test and TSWE, a test that screens students for placement into regular transfer freshman English (Eng. 131) or into one of three remedial courses (Com. 113, Com. 136, or Com. 139). At the beginning of these four courses, all students write an essay to verify the correctness of placement. They choose topics from a short list. Instructors of

[87]Helmer Myklebust, Development and Disorders of Written Language, Vol. I: Picture Story Language Test, (New York: Grune & Stratton, 1965).

[88]Myklebust in an unpublished letter to Sallie Gordon Cowgill, January 6, 1972.

five of the communications classes and of two sections of transfer English class graciously provided unmarked copies of all the essays that their students wrote on two of the topics. The more popular topic turned out to be one that matched exactly a topic that Richard Lloyd-Jones had analyzed for primary trait scoring for the National Assessment of Educational Progress.[90] The papers on that topic, the role of women, provided the base for a pilot study of ways of measuring writing quality.

I checked the TSWE scores of the writers of each paper and selected seven papers whose writers had scored 20, 25, 29, 32, 36, 44, and 58, so that the samples would represent the range of TSWE scores (20 to 60 possible). Some papers had masculine names, and one had a Spanish surname, reflecting the total college enrollment. I coded the papers, cut the names off, and assessed them anonymously by six different measures and then used the code key to find the TSWE score and the reading level of the writer of each paper. Results appear in the appendix.

When one paper seemed to be on a combination topic I rejected it temporarily. It regained its position when it turned out to be the only paper in the group on the popular topic written by someone with a TSWE score of over 45. Apparently the other topic, asking preference for the good old historical times or the present, attracted all the students with higher TSWE scores, although it was a less popular topic overall.

[90]Lloyd-Jones, op. cit.

For a holistic rating that would not influence the other analyses of these papers, two different instructors applied the College of the Mainland's English 131 "Composition Tabulation Sheet." It subtracts mechanical errors from a descendent of the Cleveland Rating Scale.[91] A copy appears in the appendix. One instructor had had ten years experience using and modifying the scale for English courses. The other instructor had just begun teaching remedial writing. Their scores were averaged. The total score represents a percentage for content and organization minus points for mechanical errors.

For primary trait scoring I used only the final scoring guide printed with Lloyd-Jones's article.[92] Trying to use the scoring guide without sample papers for practice and discussion convinced me of the need for preparation to get valid assessment. I analyzed the appeals according to the scoring guide, but tabulation of those few results would be meaningless. Discussing them with the writers, however, would have been instructive.

To get peer evaluation, I projected the papers to the lowest level remedial course and asked the students to grade them A, B, C, D, E, or F. They asked me to read aloud two of the papers of low legibility, one of which they rated high and one low. In the middle of the readings they paused to discuss criteria, but the students had little in mind. Their discussion made no discernable difference in their

[91] Follman, loc. cit.

[92] loc. cit.

grading of the last papers. None received E or F, so I tabulated them on a four-point scale. No obvious consistency appeared. The mean and standard deviation for each essay appear in the appendix.

To study syntactic complexity, I counted the mean T-unit length and also the number and proportion of words in final free modifier position. The results were higher than expectations based on narrative or expository writing norms. The assignment was argumentative, however, and argumentative writing attracts syntactic complexity.[93]

In listing intellectual processes, I noted that focus was always present as the grammatical subject of each sentence. Thus length directly influenced the count. I grouped and tabulated duplications indicated by pronouns or synonyms, but I did not analyze the focus further. Contrast and time sequence or change seemed to overlap, but I tried to follow Odell's rubric.[94] It seemed worthwhile to report separately the temporal and logical sequence, adding one half for the latter. Most essays used most of the processes, but there was a wider range in the total of subtypes of each process used, with duplicates eliminated. The specifics number all the subtypes except those in focus.

Computer programming for a small sample is impractical, but I did a count of the cohesive ties (types and items). See the appendix.

[93]See footnote 10.

[94]Odell, loc. cit.

I selected scores on thirteen variables for computer analysis of correlation, matching them first to the English 131 total score and then to the TSWE score. Since the pilot sample is so small, it is hard to conclude anything significant. Given the nature of the English 131 holistic total (which is content minus errors), part to whole correlation is expected. Reading correlated next best (.88) and then, perhaps surprisingly, cohesive types (.82). These should be analyzed further in studies with a larger number of samples. The correlation of cohesive types is significant at .03. See the appendix for a "Table of Estimates of Usefulness of Different Measures."

APPENDIX

APPENDIX

COMPOSITION TABULATION SHEET
ENGLISH 131

	YES	NO	VALUE	SCORE
I. CONTENT AND ORGANIZATION (100%)				
A. There is an appropriate Introduction.	___	___	10%	___
B. There is a stated or strongly implied thesis idea that controls the entire paper.	___	___	15%	___
C. The paper fits the assignment.	___	___	15%	___
D. Each body paragraph in the paper has a topic sentence that supports the thesis statement and moves in a logical orderly direction.	___	___	15%	___
E. Each body paragraph is adequately developed to prove its topic sentence.	___	___	15%	___
F. The paper includes a conclusion that smoothly ties the body of the paper back to the thesis.	___	___	10%	___
G. Ideas in the paper flow smoothly from one paragraph and sentence to the other and there are no extraneous sentences or points; i. e., the paper is free from digressions.	___	___	10%	___
H. Sentences show appropriate co-ordination, subordination, etc., and are varied in structure.	___	___	10%	___
II. MECHANICS				
A. Spelling errors	___	___	1 each	___
B. Punctuation errors other than those included in run-on and fragments.	___	___	1 each	___
C. Errors in the use of modifiers other than dangling modifiers	___	___	1 each	___

D. Dangling modifier	___	___	3 each	___
E. Errors in the use of pronouns i.e. agreements with antecedent, unnecessary shifts in person, case, vague antecedent.	___	___	3 each	___
F. Errors in verb tense and agreement	___	___	3 each	___
G. Fragments run-on and comma splice.	___	___	3 each	___
H. Inadequate wording (words out of order, omissions, inaccuracy in usage)	___	___	1 each	___
I. Nonparallel Structure	___	___	1 each	___
J. Dull wording (careless repetition, passive or form of to be when strong action verb is needed, etc.)	___	___	1 each	___

III. POSSIBLE POSITIVE POINTS	VALUE	SCORE
A. Outstanding vocabulary (diction)	10%	_____
B. Imaginative approach to subject	10%	_____
C. Depth of understanding of subject	10%	_____
D. Ability to analyze a problem in an unusually perceptive or original way	10%	_____

FOOTNOTE:

The maximum number of points that a student may earn on one theme is 100. Positive points are seldom earned and indicate a merit beyond what is normally found in freshman papers.

Intellectual Process Cues

(from Odell, "Measuring Changes in Intellectual Prosesses," pp. 108-120 in Evaluating Writing by Cooper and Odell)

Process	Cues
Focus	Grammatical Subject of Each Clause
Contrast	Connectors (such as or, else, but, instead, though)
	Comparative and superlative forms
	Negatives (including without)
	Nouns, verbs, adjectives, and adverbs stating contrast
Classification	Predicate nominatives
	Example and instance labels or phrases
	Words meaning similar, resemble, class, and synonyms
Change	Words meaning change and its synonyms
	Verbs that could be rewritten with become
	Verb phrases that include a synonym for begin or stop plus a verbal
Physical Context	Nouns referring to geographical location, objects in physical setting, or sensory properties of a physical setting

Sequence in Time

Adverbial element meaning before, during, or after

Logical Sequence

Words implying cause-effect relationship (because, since); the phrase if....then.

MEASUREMENT OF COHESION

DEFINITIONS AND EXAMPLES OF TYPES OF COHESION

REFERENCE

1. Personal Pronoun: Third person personal pronouns: he, him, his, she, her, hers, it, its, they, them, their, theirs.

2. Demonstratives and The: This, these, that, those, here, there, now (rarely), then, the.

3. Comparatives: Same, identical; similar (-ly), such; other, different else; more, less, as many, ordinal numbers; as & adjective, comparatives. (Superlatives are self-defining and thus not cohesive.)

SUBSTITUTION AND ELLIPSIS

1. Noun: Omission of noun or all or part of a noun group, usually represented by words such as mine, this, his, hers; any, each, many; first, next, last; one, two, more, lots; best, better, good, superlatives or comparatives. Also substitution of one or ones (accompanied by a defining modifier) or same to replace a count noun thing, nominalized process, attribute, or fact.

2. Verb: Omission of a word or words from a complete verbal group, including markers of tense, mood, or voice. Also substitution of do for the verb.

3. Clause: Omission of a clause structure, usually in response to a question that is asked, as response to a statement, or in a reporting sequence. Also substitution of so, not, yes, or no for a clause.

CONJUNCTION

1. Adding: and, also, nor, or, or else, and not, and also, furthermore, add to that, alternatively, by the way, incidentally; that is, in other words, e.g.; thus; likewise, in the same way, on the otherhand.

2. Contrasting: Yet, though, only, but, however, even so, all the same, in fact, actually; but, and, however, conversely; instead, on the contrary, at least, I mean, or rather, in any case, anyhow.

3. Cause: So, then, therefore, consequently; on account of this, in consequence, with this in mind, for, because; it follows, arising out of this, to this end; then, in that case, in such an event, under the circumstances, otherwise, under other circumstances; here, in this respect, apart from this.

4. Time & Summation: Then, next, just then, before that, hitherto, in the end; first, then, finally; at once, soon, next time, next day, meanwhile, until then, at this moment; in conclusion, up to now, at this point, from now on; to sum up, to resume.

5. Continuing: Now, of course, well, anyway, surely, after all.

LEXICAL

1. Same: Repetition of the same or closely related words, including inflections and derivations, not necessarily with the same meaning or referent.

2. Synonym: Synonyms; near synonyms that are different but related members of the same subgroup, such as Tuesday and Thursday or king and queen; hyponyms more specific than the original word; semantically related words expected to the found together and interpreted in terms of each other (tree and leaf).

3. Higher Category: Superordinates; words of a more general, higher category that includes the original item as a subset of it.

4. General Nouns: Generalized reference terms that could be replaced by a pronoun (R 1), usually accompanied by the reference item the or a demonstrative (R 2). They can carry an attitudinal meaning of adjective (the poor thing). People, person, man, woman, child, boy, girl, creature; thing, object; stuff; business, affair, matter; place; question, idea.

MEASUREMENT OF COHESION: TABULATION

IDENTIFICATION OF WRITING TO BE ANALYZED _____ DATE _____

Number of Types Found:	Cohesive Items		Referent		Other Occurences
TYPE OF COHESION	Word(s)	No. of Sentence	Word(s)	No. of Sentence	No. of Sentence
Reference					
1. Personal Pronoun					
2. Demonstratives & The					
3. Comparatives					
Substitution & Ellipsis					
1. Noun					
2. Verb					
3. Clause					
Conjunction					
1. Adding					
2. Contrasting					
3. Cause					
4. Time & Summary					
5. Continuing					
Lexical					
1. Same					
2. Synonym					
3. Higher Category					
4. General Noun					

Number of sentences analyzed:

Evaluation:

Comment:

Location: Beginning ____
Middle ____
End ____
Complete ____

TABLE OF RESULTS

Factor \ Paper	1	2	3	4	5	6	7	Mean	S.D.	Correlation With Holistic Scale Total
1. Holistic Scale Total	29	31	34	45	61	72	85	51.0	20.4	1.0
2. Student Ratings	3.1	2.3	2.1	3.0	2.9	3.4	3.0	2.8	0.4	.56
S.D.	0.9	0.9	0.7	0.6	0.8	0.8	--	--	--	--
3. Intellectual Process:										
Specifics	17	11	9	15	14	22	17	15.0	4.0	.61
Types	17	11	11	15	14	33	28	18.4	--	--
Tokens	50	23	23	57	36	47	44	40.0	--	--
4. Cohesive Types	5	5	5	4	7	6	9	5.9	1.6	.82
5. Cohesive Items	21	11	14	26	23	23	18	19.4	5.0	.35
6. Final Free Modifiers: Percent	42	21	37	40	24	32	25	31.6	7.8	-.41
Words per T-Unit	6.2	3.6	4.0	5.9	2.6	4.1	4.2	4.4	--	--
7. Length: Words	205	137	98	282	151	178	273	189.1	63.8	.43
T-Units	14	8	9	19	14	14	25	14.7	--	--
8. Mean T-Unit Length	14.7	17.1	10.9	14.8	10.8	12.7	11.0	13.1	2.3	-.58
9. Primary Trait Supports	1	1	3	4	4	3	4	2.9	1.2	.68
10. Holistic Content	44	46	48	61	72	81	97	64.1	18.6	.99
11. Reading Level	6.8	9.8	7.7	9.7	13.1	13.7	12.9	10.5	2.5	.88
12. Errors in Mechanics	14	14	16	16	12	10	12	13.4	2.1	-.74
13. TSWE	29	44	20	25	32	36	58	34.9	11.9	.61

CORRELATION

Matrix

		1	2	3	4	5
1.	Holistic Total	1.0000	.5669	.6067	.8268	.3505
2.	Student Rating	.5669	1.0000	.9544	.2651	.8075
3.	Intellectual Specifics	.6067	.9544	1.0000	.3019	.6366
4.	Cohesive Types	.8268	.2651	.3019	1.0000	-.0290
5.	Cohesive Items	.3505	.8075	.6366	-.0290	1.0000
6.	% Final Modifier	-.4074	.1935	.1300	-.5862	.4259
7.	Total Words	.4309	.6268	.5167	.2238	.5723
8.	Mean T.U. Length	-.5888	-.0794	-.0682	-.6222	-.1859
9.	P. Trait Support	.6868	.2496	.1447	.4329	.5164
10.	Holistic Content	.9967	.5601	.5898	.8257	.3533
11.	Reading Level	.8880	.4856	.5207	.6733	.3075
12.	Mechanic Errors	-.7419	-.6315	-.7346	-.6511	-.2267
13.	TSWE	.6160	.2016	.2984	.7610	-.2654

CORRELATION

Matrix

	6	7	8	9	10
1. Holistic Total	-.4074	.4309	-.5888	.6868	.9967
2. Student Rating	.1935	.6268	-.0794	.2496	.5601
3. Intellectual Specifics	.1300	.5167	-.0682	.1447	5898
4. Cohesive Types	-.5862	.2238	-.6222	.4329	.8257
5. Cohesive Items	.4259	.5723	-.1859	.5164	.3533
6. % Final Modifier	1.0000	.2127	.0634	-.0802	-.3956
7. Total Words	.2127	1.0000	.0899	.3795	.4894
8. Mean T.U. Length	.0634	.0899	1.0000	-.6996	-.5683
9. P. Trait Support	-.0802	.3795	-.6996	1.0000	.7026
10. Holistic Content	-.3956	.4894	-.5683	.7026	1.0000
11. Reading Level	-.6302	.1979	-.4006	.5867	.8673
12. Mechanic Errors	.4849	-.0472	.2921	-.1432	-.6974
13. TSWE	-.7019	.3605	.0034	.0374	.6348

CORRELATION

Matrix

		11	12	13
1.	Holistic Total	.8880	-.7419	.6160
2.	Student Rating	.4856	-.6315	.2016
3.	Intellectual Specific	.5207	-.7346	.2984
4.	Cohesive Types	.6733	-.6511	.7610
5.	Cohesive Items	.3075	-.2267	-.2654
6.	% Final Modifier	-.6302	.4849	-.7019
7.	Total Words	.1979	-.0472	.3605
8.	Mean T.U. Length	-.4006	.2921	.0034
9.	P. Trait Support	.5867	-.1432	.0374
10.	Holistic Content	.8673	-.6974	.6348
11.	Reading Level	1.0000	-.8135	.5521
12.	Mechanic Errors	-.8135	1.0000	-.5540
13.	TSWE	.5521	-.5540	1.0000

TABLE OF ESTIMATES OF USEFULNESS OF DIFFERENT MEASURES

	Administration:	Content Validity:	Placement:	Diagnosis:	Growth:	Total Points:
Holistic Current Method	2	9	7	7	8	33
Holistic Quicker Method	6	8	8	2	4	28
Primary Trait	2	6	6	6	6	26
Peer Rating	0	5	1	4	6	16
TSWE	8	4	8	2	6	28
Syntactic Complexity	2	5	5	6	7	25
Intellectual Process	2	6	3	4	7	22
Computer	2	6	4	6	6	24
Cohesion	6	7	5	2	8	28
Myklebust	2	8	8	8	8	34

MEANING OF POINTS

0	Not applicable
1	Undesirable
2	Not feasible
3	Weak
4	Questionable, variable
5	Unknown
6	Incomplete, limited, partial use
7	Possible
8	Feasible
9	Good

BIBLIOGRAPHY

Ajay, H. B. Strategies for Content Analyses of Essays by Computer. Dissertation Abstracts International 34 (1973), 2375A.

Beaven, May H. "Individualized Goal Setting, Self-Evaluation, and Peer Evaluation," Evaluating Writing. ed. Charles R. Cooper et al. Urbana, NCTE, 1977. 135-156.

Bishop, Robert L. "Computing in the Teaching of Journalistic Skills," On-Line 3 (May, 1974). no. 3, 5-12.

Boder, David P., "The Adjective-Verb Quotient: A Contribution to the Psychology of Language." Psychological Record 3 (1940). 310-343.

Bortz, Donald Richard. "The Written Language Patterns of Intermediate Grade Children When Writing Compositions in Three Forms: Descriptive, Expository, and Narrative." DAI 30 (1970), 5332 A (Lehigh).

Braddock, Richard. "Evaluation of Writing Tests." Reviews of Selected Published Tests in English. ed. Alfred H. Grommon, Richard Braddock, J. N. Hook, William A. Jenkins, Walter Loban, and Alan C. Purves, NCTE Committee to Review standardized Tests. Urbana: NCTE, 1976.

Braddock, Richard. Research in Written Composition. Champaign: NCTE, 1963.

Breland, Hunter M. A Study of College English Placement and the TSWE, Project Report 77-1. Princeton: ETS, 1976.

Breland, Hunter M. Group Comparison of the Test of Standard Written English, Research Bulletin RB-77-15. Princeton: ETS, 1977.

Britton, James et al., The Development of Writing Abilities (11-8). London: Macmillan, 1975.

Britton, James. Language and Learning. Middlesex, England: Penguin Books Ltd., 1970.

Callis, Robert, and Willoughby Johnson. Missouri College English Test Manual. New York: Harcourt, 1965. p. 14.

Christensen, Francis. Notes Toward a New Rhetoric. New York: Harper and Row, 1967.

Clark, Sandra. "Problems with the Test of Standard Written English." CLAC - 4. March, 1978, 10-13.

College Entrance Examination Board. Comparative Guidance and Placement Program: Sentences. Princeton: ETS, 1967.

Cooper, Charles R. "Holistic Evaluation of Writing." Evaluating Writing: Describing, Measuring, Judging. Ed. Charles R. Cooper and Lee Odell. Urbana: NCTE, 1977.

Cooperative Test Division. Cooperative English Tests: Technical Report. Princeton: Educational Testing Service, 1960.

Cooperative Test Division. Cooperative Sequential Tests of Educational Progress. Princeton: Educational Testing Service, 1957.

Corder, Jim W. Handbook of Current English. Glenview, Ill.: Scott, Foresman, 1978.

Dauterman, Fritz. "Syntactic Maturity Test for Narrative Writing." Measures for Research and Evaluation in the English Language Arts. William T. Fagan et al. Urbana, Ill.: ERIC and NOTE, 1975.

Diederich, Paul B. Measuring Growth in English. Urbana: NCTE, 1974.

Diederich, Paul B. "How to Measure Growth in Writing Ability." English Journal 55. April, 1966. 435-49.

Division of Curriculum and Instructional Services (Julia M. Gottesman and Marian S. Schilling). A Common Ground for Assessing Competnece in Written Expression. Downey, Calif.: Los Angeles County Superintendent of Schools, 1978.

Dixon, Edward. "Syntactic Indexes and Student Writing Performance." Elementary English 49 (May, 1972), 714-716.

Dupuis, Mary M. "Transformational Analysis of Compositions (TAC)." Measures for Research and Evaluation in the English Language Arts. Ed. William T. Fagan. Urbana: ERIC and NCTE, 1975.

Educational Skills Tests, College Edition. Monterey, Calif.: CTB/McGraw-Hill, 1971.

Educational Testing Service. Advanced Placement Course Description: English Composition and Literature. Princeton: CEEB, 1978.

Educational Testing Service. Guide to the Admissions Testing Program 1978-79. Princeton: CEEB, 1978.

Educational Testing Service. The Test of Standard Written English: A Preliminary Report. Princeton: ETS, 1975.

Elbow, Peter. Writing Without Teachers. New York: Oxford Univ. Press, 1973.

Emig, Janet. The Composing Processes of Twelfth Graders. Urbana: NCTE, 1971.

Endicott, Anthony L. "A Proposed Scale for Syntactic Complexity." RTE 7 (1973), 5-12.

Explanatory and Persuasive Letter Writing. Writing Report No. 05-W-03. Denver: National Assessment of Educational Progress, 1977.

Fagan, William T., Charles R. Cooper, and Julie M. Jenson. Measures for Research and Evaluation in the English Language Arts. Urbana: ERIC and NCTE, 1975.

Finn, Patrick J. "Computer-Aided Description of Mature Word Choices in Writing." Evaluating Writing. Ed. Charles Cooper. Urbana: NCTE, 1977.

Follman, John C. and James A. Anderson. "An Investigation of the Reliability of Five Procedures for Grading English Themes." RTE 1 (1966),190-200.

Frogner, Ellen. "Problems of Sentence Structure in Pupils' Themes." English Journal 22 (1933), 742-749.

Funkhouser, Keturah. "TSWE: 'T is Not for Me." CLAC-3, (Nov. 1977), 13-16.

GED Testing Program. Information Pamphlet on the New GED Tests. Washington: American Council on Education, 1975.

Ginsburg, Herbert, and Sylvia Opper. Piaget's Theory of Intellectual Development. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969.

Godshalk, Fred I., Frances Swineford, and William E. Coffman. The Measurement of Writing Ability. New York: Educational Testing Service, CEEB, 1966.

Hackman, Judith D. and Paula Johnson. Yale College Freshmen: How Well Do They Write? New Haven: Yale Univ., 1976.

Hackman, Judith D. and Paula Johnson. "Yale: How Well do Freshmen Write? Implications for Placement and Pedagogy." College and University, 53 (Fall, 1977) 81-99.

Halliday, Michael A. K. and Ruqaiya Hasan. Cohesion in English. London: Longman, 1976.

Hiatt, Mary P. "The Feminine Style: Theory and Fact." CCC 29 (October, 1978), 222-226.

Hirsch, E. D., Jr. The Philosophy of Composition. Chicago: University of Chicago Press, 1977.

Hooks, J. N. "Tests on the English Language." Reviews of Selected Published Tests in English. Ed. Alfred H. Grommon. Urbana: NCTE, 1976.

Hunt, Kellogg W. "Early Blooming and Late Blooming Syntactic Structures." Evaluating Writing. Ed. Charles R. Cooper. Urbana: NCTE, 1977.

Johnson, Lois V. "Children's Writing in Three Forms of Composition." Elementary English 44 (1977), 265-269.

Judine, Sister M., I. H. M. A Guide for Evaluating Student Composition. Urbana: NCTE, 1965.

Kincaid, Gerald Lloyd. "Some Factors Affecting Variations in the Quality of Students' Writing." Diss., Michigan State University, 1953. Research in Written Composition. Richard Braddock, Richard Lloyd-Jones, and Lowell Schoer. Champaign: NCTE, 1963. pp. 83-95.

Kinneavy, James. "Sentence Combining in a Comprehensive Language Framework." Sentence Combining and the Teaching of Writing. Ed. Donald Daiker, Andrew Kerek, and Max Morenberg. Akron: Univ. of Akron, forthcoming.

Kinneavy, James. A Theory of Discourse: The Aims of Discourse. Englewood Cliffs: Prentice-Hall, Inc., 1971.

Lloyd-Jones, Richard. "Primary Trait Scoring." Evaluating Writing. Ed. Charles R. Cooper et al. Urbana: NCTE, 1977.

Loban, Walter. Language Development: Kindergarten through Grade Twelve, Research Report No. 18. Urbana, NCTE: 1976.

London Association for the Teaching of English. Assessing Compositions: A Discussion Pamphlet. London: Blackie, 1965.

Martin, R. G. "A Prediction Formula for a Sample of 'Good' Writing." Dissertation Abstracts International 29. (1968), 1221A.

Mellon, John C. "Issues in the Theory and Practice of Sentence Combining: A Twenty-Year Perspective." Sentence Combining and the Teaching of Writing. Ed. Donald Daiker, Andrew Kerek, Max Morenberg. Akron: Univ. of Akron Dept. of English, forthcoming.

Mellon, John C. Transformational Sentence-Combining: A Method for Enhancing the Development of Sentactic Fluency in English Composition. Champaign: NCTE, 1969.

Mills, Helen. "Language and Composition: Three Mastery Learning Courses in One Classroom." Journal of Basic Writing. Fall, 1976, pp. 44-59.

Moffett, James A. Teaching the Universe of Discourse. Boston: Houghton-Mifflin, 1968.

Mullis, Ina V. S. The Primary Trait System for Scoring Writing Tasks. Denver: National Assessment of Educational Progress, 1975.

Myers, A. E., Carolyn McConville, and W. E. Coffman. "Simplex Structure in the Grading of Essay Tests." Educational and Psychological Measurement 26 (1966), 41-54.

Myklebust, Helmer. Development and Disorders of Written Language Vol. I: Picture Story Language Test. New York: Grune & Stratton. 1965.

Ney, James. "Notes Toward a Psycholinguistic Model of the Writing Process." RTE (1974), 159-169.

Nietzke, D. A. "The Influence of Composition Assignment Upon Grammatical Structure." DAI 32 (1972), 5476 A.

Odell, Lee. "Measuring Changes in Intellectual Processes." Evaluating Writing. Ed. Charles Cooper and Lee Odell. Urbana: NCTE, 1977.

Paschal, Hugh H. Diagnostic Test to Accompany Handbook of Current English and Workbook of Current English. Glenview, Ill.: Scott Foresman, 1978.

Pike, Kenneth L. "Beyond the Sentence." CCC 15 (1964).

Pope, Mike. "The Syntax of Fourth Graders' Narrative and Explanatory Speech." RTE 8 (1974), 219-227.

Potter, Robert R. "Sentence Structure and Prose Quality." RTE 1 (Spring, 1967), 17ff.

"Resolution No. 1," CCC, (Oct. 1978), 309.

"Resolution No. 6" and "Resolution No. 7," CCC 25, (Oct. 1974), 339

Raygor, Alton L. Ed. McGraw-Hill Basic Skills System Writing Manual. Monterey, Calif.: CTB/McGraw-Hill, 1970.

Rosen, Harold. "An Investigation of the Effects of Differentiated Writing Assignments on the Performance in English Composition of a Selected Group of 15/16 Year Old Pupils." The Development of Writing Abilities(11-18), James Britton, Harold Rosen, et al. London: Macmillan, 1975 pp. 2-3.

Sanders, Sara E. and John H. Littlefield. "Perhaps Test Essays Can Reflect Significant Improvement in Freshman Composition: Report on a Successful Attempt." RTE 9 (1975), 145.

Seegers, J. C. "Form of Discourse and Sentence Structure." Elementary English Review 10 (March, 1933), 51-54.

Slotnik, Henry B. "Toward a Theory of Computer Essay-Grading." Journal of Educational Measurements 9 (1972), 253-63.

Slotnik, Henry B. and W. Todd Rogers. "Writing Errors: Implications about Student Writers." RTE 7 (1973), 387-98.

Task Force on Measurement and Evaluation in the Study of English, Alan Purves, Chair. Common Sense and Testing in English. Urbana: NCTE, 1975.

Vygotsky, Lev Semenovich. Thought and Language. Tr. and ed. E. Hanfmann and G. Vakar, 1934. Cambridge: MIT Press, 1962.

Webb, E. Jackson. Computer Assisted Instruction Program in Writing. Olympia, Wash: Applied Linguistics, The Evergreen State College 1973.

Weltner, Klaus. The Measurement of Verbal Information in Psychology and Education. Tr. Barbara M. Crook. New York: Springer-Verlag, 1973.

Wolk, Anthony. "The Relative Importance of the Final Free Modifier: A Quantitative Analysis." RTE 4 (1970), 59ff.

Woodfin, M. J. "The Quality of Written Expression of Third Grade Children Under Different Time Limits." Journal of Experimental Education 37, No. 3 (1969), 89-91.

Writing Mechanics, 1969-1974: A Capsule Description of Changes in Writing Mechanics, Writing Report No. 05-W-01. Denver: National Assessment of Educational Progress, 1975.

Write/Rewrite: An Assessment of Revision Skills, Writing Report No. 05-W-04, 1977. Denver: National Assessment of Educational Progress, 1977.

Young, Richard E. Alton L. Becker, and Kenneth L. Pike. Rhetoric: Discovery and Change. New York: Harcourt, Brace & World, Inc., 1970 , p. 127.

UNIVERSITY OF CALIF.
LOS ANGELES

JUN 22 1979

CLEARINGHOUSE FOR
JUNIOR COLLEGES